Chapter 22

USE OF TIME-FREQUENCY REPRESENTATIONS IN THE ANALYSIS OF STOCK MARKET DATA

Gonul Turhan-Sayan
Department of Electrical and Electronics Engineering
Middle East Technical University
06531 Ankara, Turkey
gtsayan@eee.metu.edu.tr

Serdar Sayan *
Department of Economics
Bilkent University
06533 Ankara, Turkey
sayan@bilkent.edu.tr

Abstract The analysis of economic/financial time series in the frequency domain is a relatively underexplored area of the literature, particularly when the statistical properties of a time series are time-variant (evolutionary). In this case, the spectral content of the series varies as time progresses, rendering the conventional Fourier theory inadequate in describing the cyclical characteristics of the series fully. The joint Time-Frequency Representation (TFR) techniques overcome this problem, as they are capable of analyzing a given (continuous or discrete) function of time in time and frequency domains simultaneously.

To illustrate the potential of some of the TFR techniques widely used in various fields of science and engineering for use in the analysis of stock market data, the behavior of ISE-100 index of the Istanbul Stock Exchange is analyzed first, using two linear (the Gabor Transformation and the Short Time Fourier Transform) and two quadratic (the Wigner Distribution and the Page Distribution) TFRs. The performance of each TFR in detecting and decoding cycles that may be present in the original ISE data is evaluated by utilizing a specially synthesized

*The authors are truly grateful to Mr. M. Mehdi Jelassi for the skillful assistance he provided during the preparation of camera-ready copy of the manuscript.

E.J. Kontoghiorghes et al. (eds.),
Computational Methods in Decision-Making, Economics and Finance, 429–453.

time series whose trend and/or cycle components can be analytically specified and computed. This series is constructed in such a way to roughly mimic the pattern of a stock index series such as the original ISE series and is used as a benchmark for comparative performance analysis. The results indicate that the performance of the Page distribution, used for the first time in economics/finance literature, is significantly superior to the other TFRs considered. The analysis is then repeated using NASDAQ-100 index data recorded over the last 15 years so as to see if the results are robust to a change in the source of stock data from an emerging to a well-established market. The results point to a superior performance by the Page distribution once again, demonstrating the robustness of our previous results.

Keywords: Business cycles, Time-frequency representations, Stock index series, Page distribution, Wigner distribution, Gabor transformation, Short time Fourier transform.

1. Introduction

The random-walk (white-noise) process has long provided a popular tool to model the behavior of stock prices, dominating the finance literature. This popularity has been challenged by studies reporting deviations from the random walk hypothesis based on the analysis of long-run stock returns –see, for example, [Lo and MacKinlay (1988)]; [Fama and French (1988)], and [Poterba and Summers (1998)]. Explaining why such deviations occur requires formulating an alternative hypothesis.

[Chen (1996a)] suggests a mixed process with random noise and deterministic patterns, including the possibility of deterministic chaos, to explain the behavior of stock prices over time. He argues that the recognition of the existence of persistent chaotic cycles presents a new perspective on the reasons underlying market volatility, by pointing to new sources of economic uncertainties. By noting that the competition may not eliminate the nonlinear pattern in the stock market under non-equilibrium situations with an evolving trend and shifts in frequencies, he suggests what he calls the color-chaos model of stock-market movements as an alternative to the random walk approach. Through this approach, [Chen (1996a)] contends, a link between the business cycle theory and the asset-pricing theory may be established. Given that real stock prices should converge to the expected value of discounted future cash flows, real stock prices must indeed reflect the cyclicality of real output through the effects of output cycle on cash flows. Furthermore, discount rates of future cash flows are composed of term and default premia, and each of these is known to co-vary with the business cycle [Jacquier and Nanda (1991)]. If this reasoning is correct, cyclical models would be a natural alternative to the random walk.

Within this framework, stock price or return series can essentially be viewed as composed of trend and cycle components, as well as some additive noise. Furthermore, there would possibly be a number of dominant cycles oscillating

at different frequencies, and one of these cycles is expected to be the business cycle.

Using a cyclical approach in developing an empirical framework to describe the movement of stock prices or returns over time, or testing the random walk hypothesis against cyclical alternatives requires detecting the existence of business cycles and identifying their periodicity. [Sargent (1979)] offers two closely related (but not necessarily equivalent) definitions of a business cycle by considering a single series governed by a stochastic difference equation. According to the first definition, a variable is said to possess a cycle of a given frequency, if its covariogram exhibits damped oscillations at that frequency. By the second definition, a given time series has a cycle if a peak occurs in the spectral density function of the series. The spectral density function is nothing but the squared magnitude of the discrete Fourier transform of the time series. Therefore, a spectral peak observed at a specific frequency f implies the existence of a pair of spectral poles which are complex conjugate of each other and lead to a sinusoidal component in the time series oscillating at that particular frequency, i.e., implies the existence of a cycle with a period $T = \frac{1}{f}$. If this period is within the range 2 to 4 years, the cycle is called a minor business cycle, and if it falls in the range extending up to 8 years, the cycle is called a major business cycle by the NBER definitions [Sargent (1979)].

As these two definitions would indicate, a time series may be analyzed either in the time domain or in the spectral (or frequency) domain, both producing essentially the same results. Transformation of the signal representation from one domain to another is achieved by using the Fourier transform (FT) and the inverse Fourier transform (IFT) operations. If the spectral content of a given signal varies as time progresses, however, the conventional Fourier theory fails to fully describe the contribution of arbitrarily chosen spectral components over certain time bands. The time-frequency representation (TFR) techniques have emerged as viable solutions to this challenging problem, since they make it possible to analyze a given (continuous or discrete) function of time in time domain and frequency domain *simultaneously*. Providing localization both in time and frequency (within the resolution limits allowed by the uncertainty principle), TFRs can describe the variation of a function in the two-dimensional joint time-frequency domain in detail. The Gabor transform (GT), the short time Fourier transform (STFT) and the Wavelet transform (WT) are *linear* time-frequency representations, whereas the Wigner distribution (WD), the Spectrogram (magnitude square of STFT), the Scalogram (magnitude square of the WT), the Choi-Williams distribution (CWD) and the Page distribution (PD) are some of the well-known *quadratic* time-frequency representations –see [Hlawatsch and Boudreaux-Bartels (1992)].

The purpose of this chapter is to illustrate the use of some of these TFR techniques in the analysis of stock price data from emerging as well as well-

established stock markets, and to compare their performances. For this purpose, we first analyze the behavior of ISE-100 index of the Istanbul Stock Exchange over the period from July 9, 1990 to December 25, 2000 using two linear TFRs (the GT and the STFT) and two quadratic TFRs (the WD and the PD). In order to evaluate the comparative performance of these TFRs in detecting and decoding cycles that may be present in the original ISE data, we utilize a specially synthesized time series whose trend and/or cycle components can be analytically specified and computed. We construct this series in such a way to mimic the pattern of a stock index series such as the ISE-100 and let it serve as a benchmark in our simulated performance analysis. We then test the robustness of results to a switch from an emerging to a well-established (yet, relatively volatile) stock market by using an alternative series containing the NASDAQ-100 index values recorded over the last 15 years.

The analysis of economic/financial time series in the frequency domain is a relatively underexplored area of the literature. Examples include [King and Rebelo (1993)], [Bowden and Martin (1993)], [Thoma (1994)], [Levy and Chen (1994)], [Garcia-Ferrer and Queralt (1998)], [Hong (1999)], and [Bjornland (2000)]. The studies using TFRs are even fewer in number and are based mostly on linear TFRs such as the wavelet transform –see, [Greenblatt (1997)], and [Lee (1998)], for a review article– and the Gabor transform, with the Wigner distribution being the only quadratic TFR employed –see [Chen (1996a)], and [Chen (1996b)]. Thus, the present study is expected to be a significant contribution to the literature not only because one of the quadratic TFRs (i.e., the PD) is used for the first time in the finance literature in general, but also because the application of all four TFRs to financial data from an emerging stock market is new to this study. Furthermore, the study develops a novel approach for evaluating the comparative performances of the TFR techniques considered by constructing a synthesized series with known analytical features to serve as a benchmark, and further increases the usefulness of this benchmark through the addition of some Gaussian noise. Finally, the study deserves special attention also because the results obtained highlight a significant potential for the PD launched into economics/finance literature here as a useful tool for future research in these fields.

The rest of the discussion in the paper is organized as follows. The next section briefly describes the theoretical framework employed. Section 3 explains the implementation of TFR analysis and reports results from each of the four TFR techniques considered for each of the simulated and true stock series. Section 4 discusses the robustness of the results by repeating the analysis using NASDAQ data, and Section 5 concludes the paper.

2. The theoretical framework

Fourier transform theory states that a given time series can equivalently be characterized either in time domain or in frequency domain. In general, transformation of the signal representation between the time domain and the frequency domain (also known as the spectral domain) is achieved by computing the Fourier transform (FT) and the inverse Fourier transform (IFT) as given in Equations (7.1) and (7.2), respectively

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt \tag{22.1}$$

and

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft}df, \tag{22.2}$$

where $x(t)$ is the time function evaluated at time $t, X(f)$ is the Fourier transform evaluated at frequency f, and $j = \sqrt{-1}$ is the unit imaginary number. As implied by Equation (7.2), a given signal $x(t)$ can be linearly decomposed into a basis of complex exponential functions, $e^{j2\pi ft} = \cos(2\pi ft) + j\sin(2\pi ft)$, oscillating at different frequencies. Then, the related decomposition coefficients $X(f)$ are computed as a function of frequency by Equation (7.1) to form the Fourier transform in the spectral domain. In other words, the magnitude of the complex valued function $X(f)$ at a given frequency f represents the strength of the signal's spectral component oscillating at that specific frequency.

For discrete-time problems, the discrete Fourier transform (DFT) and the inverse discrete Fourier transform (IDFT) need to be used instead of the continuous-time FT and IFT pair given in Equations (7.1) and (7.2). The Fast Fourier Transform algorithm (FFT) and its inverse (IFFT) are computationally optimized signal processing tools that can be used to compute Fourier transform pairs for discrete signals. While the characterization of discrete data in time domain requires the techniques for the analysis of time series, characterization in frequency domain calls for the techniques of discrete spectral analysis, both producing essentially the same results. For a given application, the choice between the time and frequency domains depends on the relative simplicity of the specific techniques available for the solution.

Both continuous and discrete Fourier transforms have proved indispensable as data analysis tools for stationary signals. Yet, if the statistical properties of a time signal are time-variant and hence, its spectral content varies as time progresses, the conventional Fourier theory becomes inadequate to fully describe the signal characteristics. Fortunately, the TFR techniques are available to overcome this problem. The TFRs are capable of analyzing a given (continuous or discrete) function of time in time and frequency domains simultaneously. In other words, they can characterize a given time signal in the two-dimensional joint time-frequency domain enabling localization both in time and frequency

within the resolution limits allowed by the uncertainty principle [Cohen (1995)]. As such, the TFRs in general may be viewed as a nonparametric approach for generalized spectral analysis of the evolutionary time series [Qian and Chen (1996)].

We now proceed by briefly describing the TFRs that we employ to identify the cycles in our simulation and test problems: GT, the Gabor transform (named after Gabor, a Nobel laureate physicist); STFT, the short time Fourier transform; WD, the Wigner distribution (developed by Wigner, another Nobel laureate in physics), and PD, the Page distribution. The first two of these TFRs are linear transforms whereas the latter two are nonlinear. (More information about the theoretical foundations of these TFRs can be found in works cited in the extensive survey article by [Hlawatsch and Boudreaux-Bartels (1992)], which also includes numerous aplications of the TFRs in science and engineering.) The choice of the set of TFRs used here out of a larger set of similar representations available in the literature has been motivated by their performance at early stages of this research, as well as in an electromagnetic target identification problem previously studied by the senior author –see [Turhan-Sayan and Karaduman (2001)]. The inclusion of GT and WD has been further motivated by their use in a previous study by [Chen (1996a)] in a context similar to ours. Due to the differences in their cross-term structures, the PD has proved to be much more useful in frequency localization than the WD in the present as well as the previous electromagnetic applications, motivating us to include it as a quadratic TFR comparable to the WD. The decision to include the STFT as a linear TFR comparable to the GT has been made on similar grounds. To the best of the authors' knowledge, this is the first time that the PD is used in economics/finance literature.

The mathematical definitions of GT, STFT, WD and PD are as follows:

1 The Gabor Transform (GT):

The Gabor expansion coefficients $G_x(n,k)$ of a given time signal $x(t)$ are implicitly defined by

$$x(t) = \sum_n \sum_k G_x(n,k) g_{nk}(t) \tag{22.3}$$

with

$$g_{nk}(t) = g(t - nT) e^{j2\pi(kF)t} \tag{22.4}$$

being the basis functions of the expansion. Gabor used these time-frequency shifted Gaussian functions in his seminal 1946 study, as they are well concentrated both in time and frequency domain. The expansion coefficient $G_x(n,k)$ is expected to indicate the signal's time and frequency content around the point (nT, kF) in the joint time-frequency domain, where T and F are the time step and the frequency step, respectively –see [Hlawatsch and Boudreaux-Bartels (1992)]. A detailed

discussion on the computation of discrete Gabor transform can also be found in [Qian and Chen (1993)].

2 The Short Time Fourier Transform (STFT):

The STFT of a given time signal $x(t)$ is computed by

$$STFT_x^{(\gamma)}(t,f) = \int_{t'} x(t')\gamma^\star(t'-t)e^{-j2\pi f t'}dt', \tag{22.5}$$

where $\gamma(t'-t)$ is the chosen window of analysis which is centered at $t'=t$ and the superscript * denotes complex conjugation –see [Hlawatsch and Boudreaux-Bartels (1992)]. As implied by this definition, the STFT of a signal may be interpreted as the *local Fourier transform* of the signal around time t.

3 The Wigner Distribution (WD):

The auto-Wigner distribution of a given time signal $x(t)$ is given by

$$W_x(t,f) = \int_\tau x(t+\frac{\tau}{2})x^\star(t-\frac{\tau}{2})e^{-j2\pi f\tau}d\tau. \tag{22.6}$$

The WD is a real-valued quadratic TFR preserving time shifts and frequency shifts of the signal. The frequency (time) integral of the WD corresponds to the signal's instantaneous power (spectral energy density) as the WD satisfies the so-called *marginals*. As a matter of fact, the WD is the only quadratic TFR satisfying all of the desired properties of the *energetic* time-frequency representations [Hlawatsch and Boudreaux-Bartels (1992)] that makes it a very popular signal processing tool, despite the severe cross-term problems encountered in applications.

4 The Page Distribution (PD):

The Page distribution of a given time signal $x(t)$ is defined as

$$PD_x(t,f) = \frac{d}{dt}\left|\int_{-\infty}^{t} x(t')e^{-j2\pi f t'}dt'\right|^2. \tag{22.7}$$

The PD is also an energetic, shift-invariant, quadratic TFR like the WD. Most of the desirable properties satisfied by the WD are also satisfied by the PD except for a few, such as the lack of a finite frequency support that the WD has [Hlawatsch and Boudreaux-Bartels (1992)].

3. Implementation and results

In this section, we describe the implementation of TFR techniques we employed for identifying the cycles in the ISE-100 data we used, and present results from the analysis with the original ISE data as well as the synthesized series used as a benchmark.

3.1 Description of the ISE data used

We have worked with two different time series, one actual and one synthesized. The first series has values of the ISE-100 index (constructed from the prices of top 100 stocks traded in the Istanbul Stock Exchange) over the period from July 9, 1990 to December 25, 2000. This is the series that is of primary interest to us and was obtained from Datastream. While the length of the original daily series is 2644 days excluding the weekends and the official holidays, a shorter version with weekly values of the index was used in the analysis to speed up the TFR computations and to reduce computer memory requirements. With 512 weekly observations, this smaller version corresponds to about 9.85 years and is smoother as compared to the original series, since the data reduction process roughly mimics a low-pass filter action to remove very high frequency signal components to some extent. The length of the series was intentionally chosen to be $512 = 2^9$ (a power of 2) so as to further accelerate the computations of TFR algorithms used.

The weekly series shown in panel (a) of Figure 22.1 was presumably composed of trend and cycle components as well as some additive noise, possibly containing a number of dominant cycles oscillating at different frequencies. The purpose of TFRs used was, therefore, to detect the existence of such cycles and investigate if any of the cycles was in the business cycle range. While the NBER's criteria for minor and major cycles in the U.S. economy ranged from 2 to 8 years, our prior expectation was to observe a business cycle with a significantly shorter period than 8 years, on account of the dynamic nature of Turkish economy and high volatility observed in the ISE. Despite the lack of evidence in the literature and a consensus (or convention similar to the NBER's) on the length of business cycles in Turkish economy –see [Metin-Ozcan, Voyvoda and Yeldan (2001)], and [Alper (2000)], we expected to observe a major cycle of 2 to 4 years.

The second time series we used was a specially synthesized series with trend and/or cycle components that can be analytically specified and computed. Our purpose in constructing this series was to check the reliability of TFR results so that cycles identified by TFR techniques can be verified without any doubt attributable to the unknown characteristics of the data. The trend of this synthesized time series was chosen to be the same as the trend estimated from the logarithmic ISE-100 weekly time series so as to maximize the structural similarity between the original data and the synthesized data –see panel (b) of Figure 22.1. In estimating the trend, we used the popular Hodrick-Prescott (HP) filter – [Hodrick and Prescott (1997)]. Despite criticism raised against the HP filter (by [King and Rebelo, 1993], for example), we retained it, as our experimentation with a fifth degree polynomial fit as an alternative detrending technique produced quite similar results –see also [Turhan-Sayan and Sayan

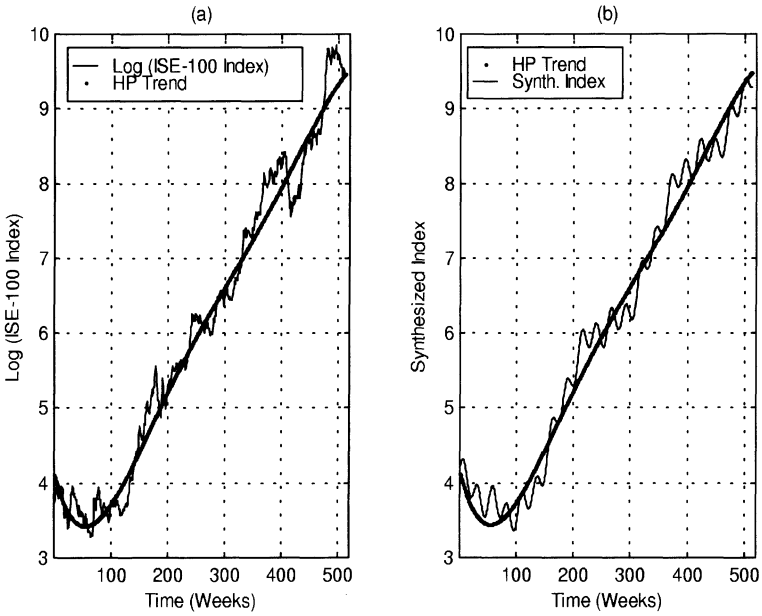

Figure 22.1. Logarithmic ISE-100 Index (a) and the Synthesized Index (b) plotted together with the trend of the Log (ISE-100) Index estimated by the HP filter.

(2001)]. [Alper (2000)] reports a similar experience with several alternatives to HP and cites other studies where the HP filter performed as well as the alternatives considered.

Once the trend was obtained using the HP filter, first, a sinusoidal cycle component was added to this trend with a period of 170 weeks (one-third of the length of ISE-100 series or approximately 3.27 years) which falls in our expected business cycle range of 2 to 4 years. Then, two more cycles were added with periods of 52 weeks (1 year) and 26 weeks (6 months) to see the effects of having more than one cycle in TFR applications. All of the cycle components were inserted to last over the entire sampling period of 512 weeks, without any damping in time. Different sinusoidal peak values of 0.3, 0.09 and 0.18 were assigned to the first, second and third cycles, respectively, to investigate the masking effect of the strong cycles on the weak cycles, if any. These peak values were chosen in such a way that not only the trends of the ISE-100 and the synthesized time series would be the same but also the strengths of their overall cycle components would be comparable. The effect of noise on the results of TFR analysis was also investigated by adding white Gaussian noise to the synthesized time series as discussed in the next subsection.

3.2 TFR results from the detrended series

Each of the time series described above contains relatively weak cycles superimposed on a strong trend component so that the trend itself is responsible for most of the total signal energy. Therefore, contribution of the cycle components to the resulting TFR outputs is obscured, unless the time series is detrended prior to the TFR analysis. To address this problem, the logarithmic ISE-100 weekly index was detrended by using the HP filter before going ahead with the TFR analysis. The following minimization problem was solved for this purpose, by taking $\lambda = 128000$:

$$\min_{\{\tau\}_{n=1}^{N}} \sum_{n=1}^{N} [x_n - \tau_n]^2 + \lambda \sum_{n=2}^{N-1} [(\tau_{n+1} - \tau_n) - (\tau_n - \tau_{n-1})]^2, \tag{22.8}$$

where the series $\{x\}_{n=1}^{N}$ represents the ISE-100 index of length $N = 512$ and the series $\{\tau\}_{n=1}^{N}$ is the trend to be derived. The resulting trend is plotted in both panels of Figure 22.1, since it is also used as the trend of the synthesized index. The overall cycle component of the logarithmic ISE-100 index was then computed as the difference between the index itself and the trend. The same procedure was also repeated to obtain the overall cycle for the synthesized index, which is obviously a composite cycle term made up of three different cycles with periods of 3.27, 1 and 0.5 years, as described earlier. The overall cycle terms obtained for the ISE-100 index and for the synthesized index are plotted in Figure 22.2.

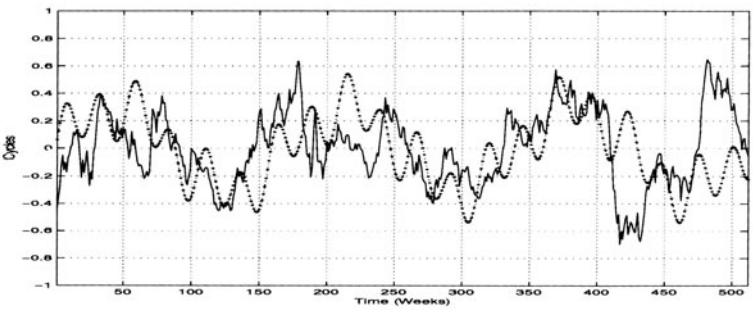

Figure 22.2. Cycles for the Logarithmic ISE-100 Index (—-) and for the Synthesized Index (....), estimated by HP detrending approach.

The final step in the suggested cycle identification process was to compute the TFRs of the overall cycle term for a given index and to study the resulting TFR output matrices to identify the individual cycle frequencies. The GT, STFT, WD and PD of detrended series were computed in MATLAB, starting with the synthesized series. This series was used as a benchmark since it is free of noise

and more importantly, we already know the cycle frequencies to be estimated for this synthesized index.

3.2.1 TFR results from the synthesized series. The contour plots for the TFR output matrices for the composite cycle term of the synthesized index as obtained from PD, WD, STFT and GT are respectively given in parts (a) to (d) of Figure 22.3 which is followed by a discussion about the performance of each TFR considered in this study.

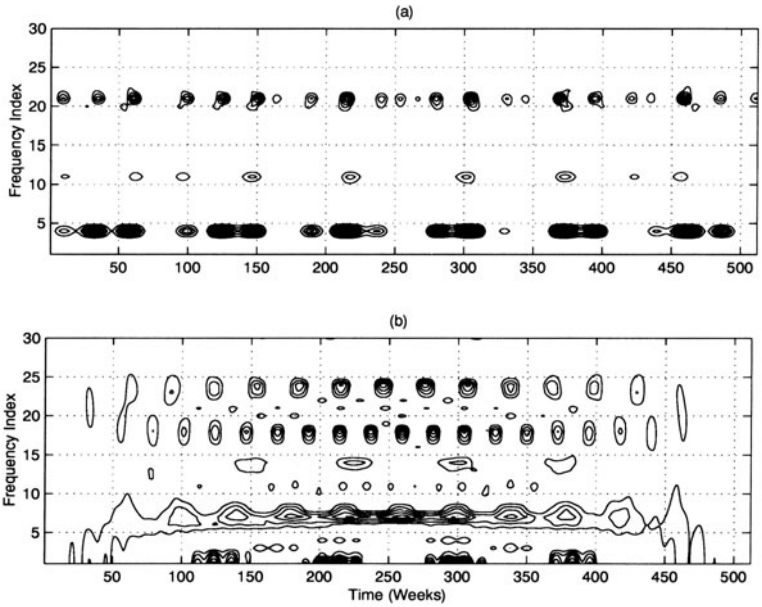

Figure 22.3. TFR Results from the Synthesized Series using (a) the PD Analysis (b) the WD Analysis.

1 Results obtained using the PD:

The PD proved to be very successful in identifying all three cycles with true periods of 170, 52 and 26 weeks by yielding the following estimates for the corresponding periods –see Figure 22.3 (a):

$$T_1 = \frac{1}{f_1} = \frac{1}{(4-1)\Delta f} \cong 170.67 \text{ weeks},$$

$$T_2 = \frac{1}{f_2} = \frac{1}{(11-1)\Delta f} = 51.2 \text{ weeks}$$

and

$$T_3 = \frac{1}{f_3} = \frac{1}{(21-1)\Delta f} = 25.6 \text{ weeks},$$

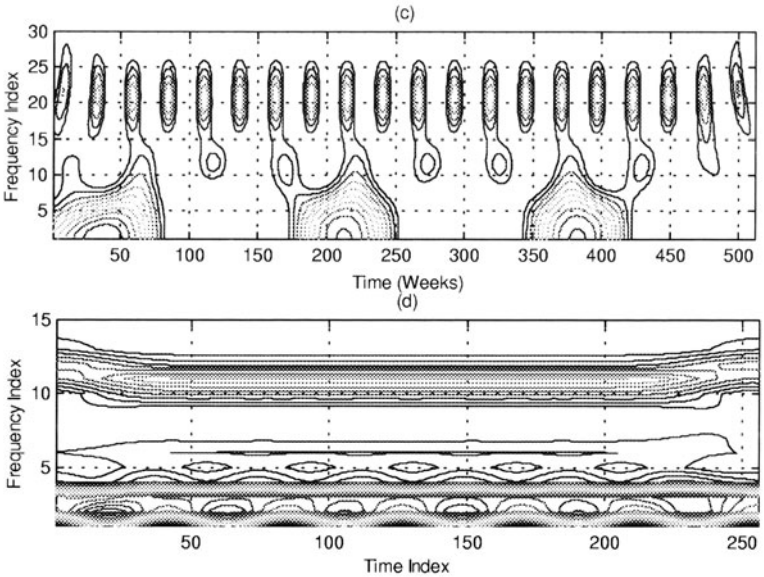

Figure 22.3. (Continued) TFR Results from the Synthesized Series using (c) the STFT Analysis (d) the GT Analysis.

where

$$\Delta f = \frac{1}{\text{Total Sampling Period}} = \frac{1}{512 \text{ weeks}}$$

is the frequency step for the vertical axis in TFR output plots, which determines the frequency resolution of the TFR computations. The time resolution, on the other hand, is determined by the time step $\Delta t = 1$ week as indicated in the horizontal axis of output plots.

The first cycle component with the largest period of 170.67 weeks (or approximately 3.28 years) can be considered as the *business cycle* for this synthesized time series. This cycle can be easily identified in the PD output contour plot shown in Figure 22.3 (a) due to the strong energy terms horizontally lined up around the frequency index level of $n = 4$. Then, the actual frequency corresponding to this level can be computed by multiplying the frequency step Δf by the number of intervals $(4 - 1)$ up to this activation level. The cycle period is simply the inverse of the cycle frequency. Hence, as compared to the true period of 170 weeks imposed while constructing the synthesized index, the estimated business cycle period of 170.67 weeks represents a computational error of about 0.4 percent. This negligible error stems from the relatively large frequency step value used in the TFR calculations due to a short total data

sampling period of 512 weeks (or 9.85 years approximately). This error in the synthesized time series problem could have been reduced further by simply increasing the time horizon to 20 years or more. However, given that the synthetic time series was designed to roughly mimic the behavior over time of the ISE-100 index, for which the data is available for only about 10 years, improving the frequency resolution was not possible for this application.

The second and third cycles of the synthesized index can also be identified similarly using Figure 22.3 (a). The horizontally extending narrow activation band lined up around the frequency index level of $n = 11$ corresponds to the second cycle with a period of 51.2 weeks (estimated with an error of 1.5 percent relative to the true period of 52 weeks). Also, another horizontally extending, moderately strong activation band lined up around the frequency index level of $n = 21$ identifies the third cycle with a period of 25.6 weeks (estimated with an error of 1.5 percent relative to the true period of 26 weeks). The activation band in the middle turns out to be weaker compared to the other two bands, as the second cycle was designed to have the lowest cycle strength to begin with.

2 Results obtained using the WD:

The WD contour plot shown in Figure 22.3 (b) turns out to be dominated by the interaction terms (WD cross terms) which are especially strong around the frequency index levels of $n = 1$ (corresponding to zero actual frequency), $n = 7$, $n = 17$, and $n = 24$. The second and the third cycles are hardly noticed at the index levels of 11 and 21 while the business cycle is severely camouflaged by very strong cross terms around. Therefore, the cycle identification results using the WD in this simulation problem are very poor as compared to those obtained using the PD, and may be deceptive.

3 Results obtained using the STFT:

The STFT contour plot shown in Figure 22.3 (c) identifies the third cycle centered at the level of $n = 21$ with a very poor resolution, and it catches the second cycle around $n = 11$ with only a very low level of certainty. It is not possible to say anything for the period or even for the presence of the business cycle, however, due to very poor frequency localization exhibited by the STFT output. In short, the performance of the STFT analysis is found to be quite poor as compared to that of the PD analysis.

4 Results obtained using the GT:

As the last test case in this problem, the GT coefficient matrix was computed by using a time-frequency sampling grid of size 256×256 with a

degree of oversampling $Q = 128$. The resulting coefficient matrix is plotted in Figure 22.3 (d). In comparing the results presented in this figure to those presented in Figures 22.3 (a-c), it should be noted that $\Delta t' = 2\Delta t = 2$ weeks and $\Delta f' = 2\Delta f = \frac{2}{512} = \frac{1}{256 \text{ weeks}}$ in the Gabor transform application. Based on this, the GT is found to be accurate in identifying the third cycle at the frequency index level of 11 –corresponding to an estimated period of $\frac{256}{11-1} = 25.6$ weeks, but with a small confidence due to poor frequency localization. Also, another strong activation band is observed as having its peak values around the frequency index level of 3 corresponding to a cycle period $T = \frac{1}{f} = \frac{1}{(3-1)\Delta f'} = \frac{1}{2/256} = 128$ weeks $\cong 2.46$ years which is far from the expected business cycle of 170 weeks in this problem. Again, the poor frequency localization does not allow for any meaningful interpretations especially for the first and second cycles.

To summarize, the WD results mainly suffer from the existence of cross terms, while the STFT and GT results exhibit serious problems in frequency localization. Therefore, the WD, STFT and GT can identify only one or two cycles with little confidence, whereas the PD successfully identifies all three cycles in the time series quite accurately. The performance of the PD remains quite satisfactory even when a realistic amount of white Gaussian noise is added to the synthesized index (Figure 22.4). The synthesized noisy index cycles are plotted in Figure 22.4 (a) while the PD contour plot computed for these cycles is given in Figure 22.4 (b). When we compare the PD analysis results for the noise-free and noisy time series –given in Figures 22.3 (a) and 22.4 (b), respectively– the weakest cycle with a period of 52 weeks is observed to be mostly affected by the added noise, as expected.

3.2.2 TFR results from the ISE-100 series. In this section, we briefly discuss the TFR results obtained from the ISE-100 index cycle. The contour plots obtained with the PD, WD, STFT and GT are given in Figures 22.5 (a) through 22.5 (d), respectively.

1 Results obtained using the PD:

Consistently with the results of the previous simulation example, the performance of the PD in detecting the cycles in ISE-100 series was found to be superior to other TFRs considered. As observed in Figure 22.5 (a), there exists a horizontal band of strong activation lined up between the frequency index levels of 4 and 6, indicating the presence of a *business cycle* with an estimated period of

$$T_1 = \frac{1}{f_1} = \frac{1}{(5-1)/512} = 128 \text{ weeks} \cong 2.46 \text{ years}$$

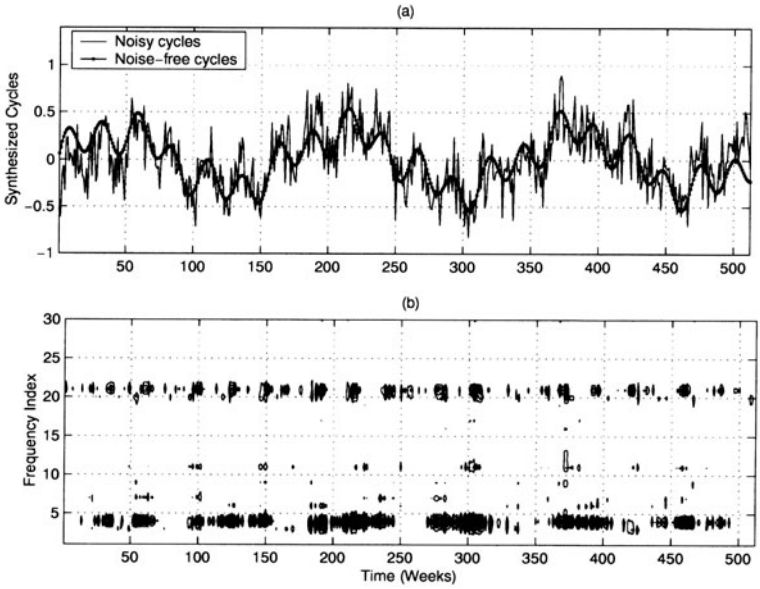

Figure 22.4. PD results for the Synthesized Series with Noise: Noisy Cycle in (a) PD Output in (b).

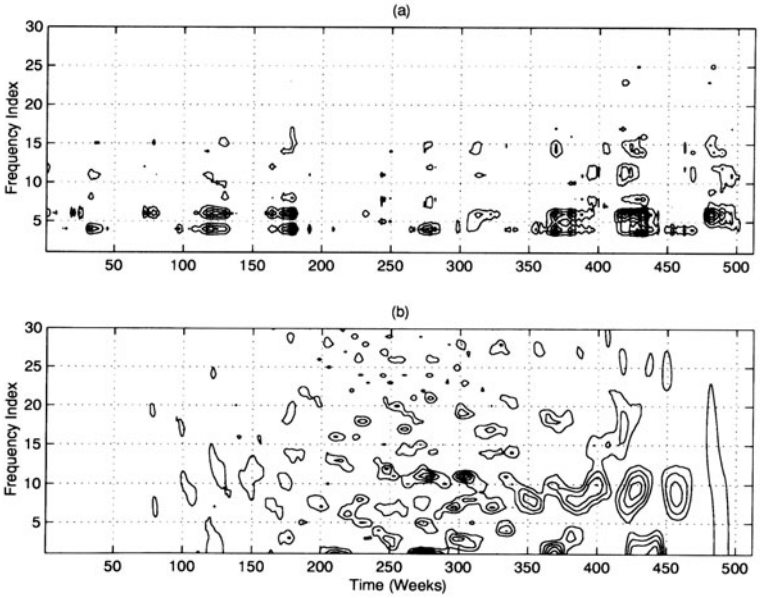

Figure 22.5. TFR Results from the ISE-100 Index Series using (a) the PD Analysis (b) the WD Analysis.

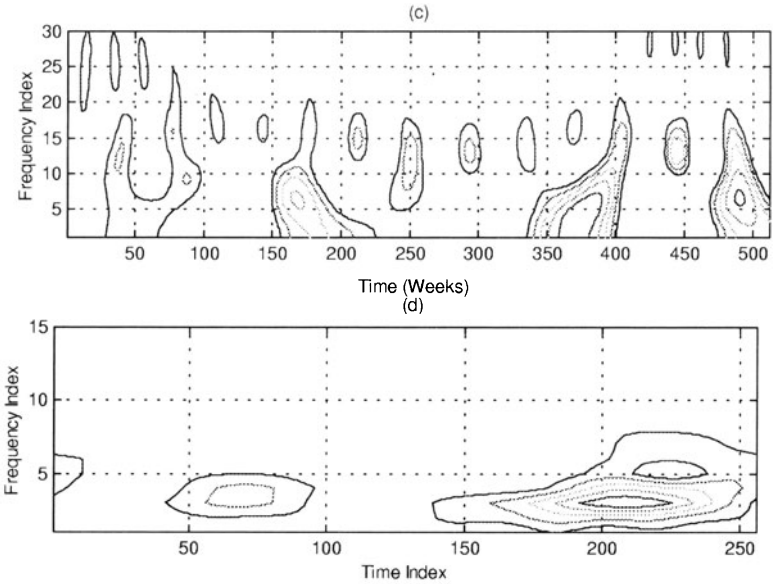

Figure 22.5. (Continued) TFR Results from the ISE-100 Index Series using (c) the STFT Analysis (d) the GT Analysis.

on the average. To be on the safe side, the period of the cycle may be said to fall within the range of $\left[\frac{1}{(6-1)/512}, \frac{1}{(4-1)/512}\right]$ weeks or approximately $[2, 3.3]$ years. Furthermore, the PD output plot indicates the existence of two more but weaker cycles at the frequency index levels of 11 and 15. The cycle periods corresponding to those frequency index levels are

$$T_2 = \frac{1}{f_2} = \frac{1}{(15-1)/512} \cong 36.6 \text{ weeks} \cong 0.7 \text{ years}$$

and

$$T_3 = \frac{1}{f_3} = \frac{1}{(11-1)/512} \cong 51.2 \text{ weeks} \cong 1 \text{ year},$$

respectively.

The actual stock data used in this application possibly contain various cycle components in varying strengths. Furthermore, as observed in Figure 22.2, the overall cycle amplitude does not remain constant but increases especially around the 150th and after the 350th weeks. In other words, the individual cycles contained in the ISE-100 index do not follow perfectly sinusoidal patterns, and hence, we do not get perfectly and uniformly concentrated activation bands in Figure 22.5 (a). The

uncertainty caused by such factors should be expected to get worse in the identification of weak cycle components.

2 Results obtained using the WD:

The WD contour plot obtained for the ISE-100 index cycles is shown in Figure 22.5 (b) which gives no indication of any cycle at all. Similar to the case encountered in the first simulation example, strong cross terms again dominate the WD output, masking all actual cycle patterns.

3 Results obtained using the STFT:

The STFT contour plot shown in Figure 22.5 (c) vaguely indicates the existence of only one cycle at the frequency index level of 6 (on the average), corresponding to a cycle period of $T = \frac{1}{f} = \frac{1}{(6-1)/512} = 102.4$ weeks $\cong 1.97$ years. As seen in the figure, the frequency resolution of the output is extremely poor. It is quite likely that this cycle (with an estimated period of about 2 years) is actually the business cycle identified earlier by the PD analysis but the period estimation of the STFT is off the mark with a margin of about 20 percent. Perhaps more importantly, the lack of certainty turns out to be the major drawback in the case of the results from the STFT analysis.

4 Results obtained using the GT:

The GT coefficient matrix for the ISE-100 index is again computed by using a time-frequency sampling grid of size 256×256 with a degree of oversampling $Q = 128$. The resulting coefficient matrix is plotted in Figure 22.5 d with the time step of $\Delta t' = 2\Delta t = 2$ weeks and the frequency step of $\Delta f' = 2\Delta f = \frac{2}{512} = \frac{1}{256 \text{ weeks}}$ as in the previous GT simulation. Based on this GT output plot, the existence of a business cycle is vaguely implied between the frequency index levels of 2 and 5 with a poor frequency resolution. The period of the suggested cycle is

$$T = \frac{1}{f} = \frac{1}{(3-1)\Delta f'} = \frac{1}{2/256} = 128 \text{ weeks} \cong 2.46 \text{ years}$$

on the average, and is the same business cycle period identified by the PD analysis. Similarly to the STFT analysis results, however, the major drawback of the results here seem to be the uncertainty in frequency localization.

4. Robustness of the results

Section 3 demonstrated the use of various TFR techniques in the identification of cycles of varying frequencies within stock market time series by using data from the Istanbul Stock Exchange, an emerging market. Of the TFR

techniques considered, the Page distribution turned out to be particularly effective in estimating the period of the business cycle within the range of 2 to 3.3 years with a good frequency localization over the entire period of investigation. The results obtained from the ISE-100 series were crosschecked against those obtained from the time series we specially synthesized.

With the aim of testing the robustness of results and evaluating the performance of the techniques considered within the context of a well-established stock market, we now consider an alternative series containing 3999 observations on the NASDAQ-100 index in this section. Originally recorded on a daily basis (excluding weekends and national holidays), the index values under consideration cover the period from November 14, 1985 to March 13, 2001, spanning about 15.3 years. This series is plotted in Figure 22.6 together with the ISE-100 data, where each time series is normalized by its own maximum value (so as not to allow either index to have values greater than 1). As can be observed from the figure, the NASDAQ series is also quite volatile, displaying sharp ups and downs especially during the last few years of the recording period.

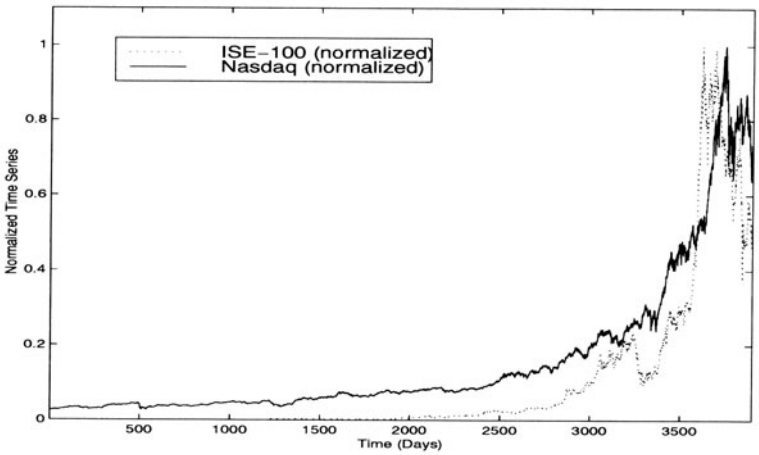

Figure 22.6. Normalized NASDAQ-100 and ISE-100 time series.

Similarly to our treatment of the ISE-100 data, we re-sampled the original NASDAQ series to work with 800 weekly observations. As the first step of the TFR analysis, natural logarithm of the NASDAQ series was detrended using an eighth order polynomial fit. The logarithmic time series and its computed trend are plotted in Figure 22.7, and the detrended series (i.e., the overall cycle component) is plotted in Figure 22.8.

Next, the PD, WD, STFT and GT output matrices were computed for the overall NASDAQ cycles, yielding the contour plots given in Figures 22.9 (a)

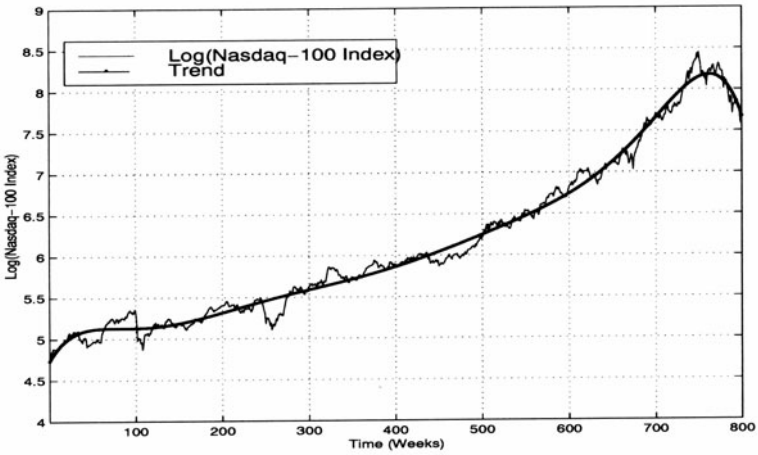

Figure 22.7. Logarithmic NASDAQ-100 Index and its Trend.

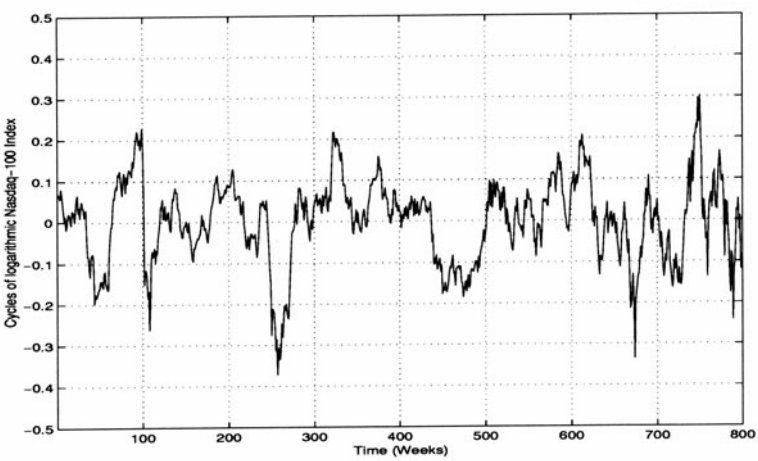

Figure 22.8. Cycle for the Logarithmic NASDAQ-100 Index.

through 22.9 (d), respectively. The NASDAQ-100 time series is resampled one more time with $\Delta t = 2$ weeks to reduce the number of observations to 400 so as to work with smaller matrices, and save on processing time required by TFR output computations. As in the case of the results of TFR analysis carried out for the ISE-100 index in Section 3, the PD signal analysis produced the best results for the NASDAQ index.

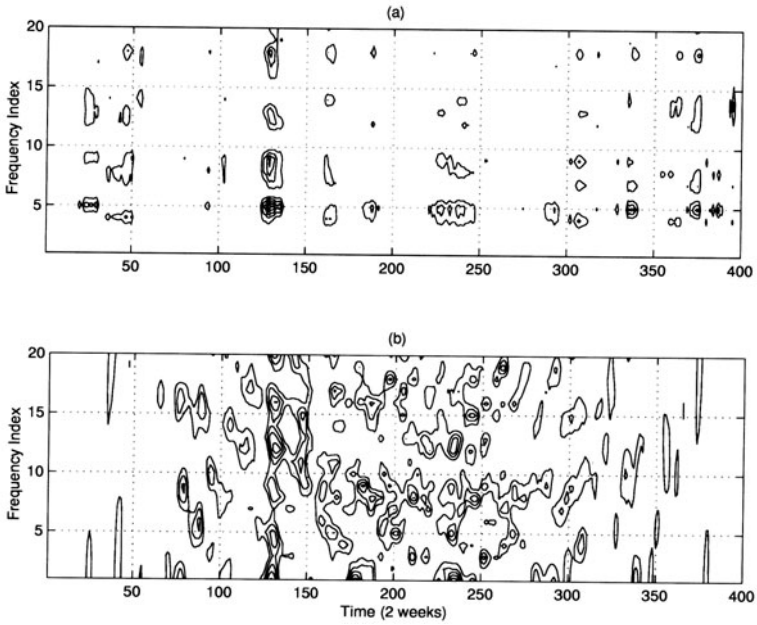

Figure 22.9. TFR Results from the NASDAQ-100 Index Series using (a) the PD Analysis (b) the WD Analysis.

Figure 22.9 (a) shows that there exists a well-concentrated signal activation band horizontally lined up around the frequency index level of 5, indicating the presence of a business cycle with an estimated period of

$$T_1 = \frac{1}{f_1} = \frac{1}{(5-1)/800} = 200 \text{ weeks} \cong 3.85 \text{ years}$$

since the frequency step, Δf, is now $\frac{1}{800}$, where 800 is the number of weeks in the total recording period. The PD output matrix plotted in Figure 22.9 (a) also indicates the presence of several other signal activation bands. While some of these are relatively strong, they are all lined up at frequency levels higher than 5, corresponding to cycle periods shorter than 2 years. Thus, the activation band around $n = 5$ is the only important band to be used in the estimation of the business cycle period.

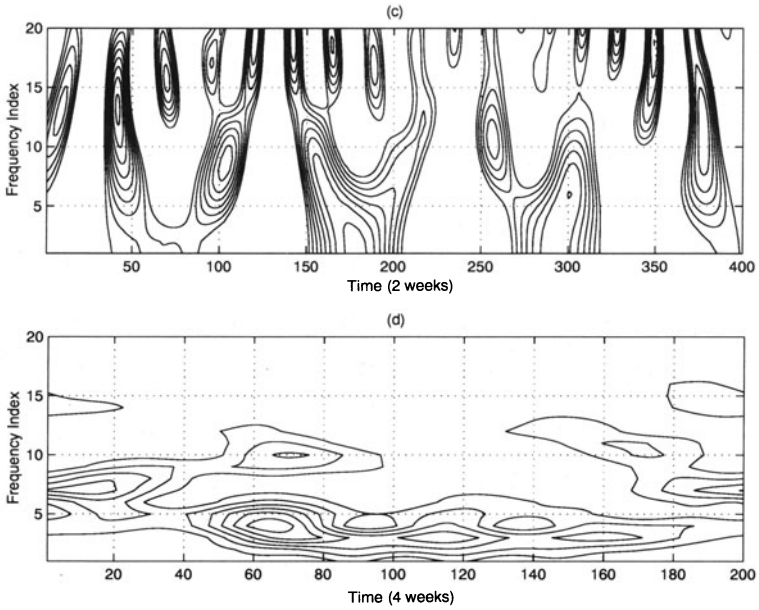

Figure 22.9. (Continued) TFR Results from the NASDAQ-100 Index Series using (c) the STFT Analysis (d) the GT Analysis.

In this application, the WD analysis did not produce any useful results to identify the business cycle, as the WD output matrix plotted in Figure 22.9 (b) is found to be suffering from the masking effect of strong cross terms. Likewise, the STFT contour plot given in Figure 22.9 (c) does not give any clear information about the business cycle due to the extremely poor frequency localizations.

The GT results displayed in Figure 22.9 (d) are computed with a 200×200 grid using a degree of oversampling $Q = 100$. We therefore have $\Delta t' = 4$ weeks and $\Delta f' = \frac{1}{400 \text{ weeks}}$ in this figure. The signal activation band observed between frequency index levels of 1 and 6 indicates the presence of the business cycle. Taking $n = 3$ as the average level of this band, the period of the business cycle can be roughly estimated as

$$T_1 = \frac{1}{f_1} = \frac{1}{(3-1)/400} = 200 \text{ weeks} \cong 3.85 \text{ years}$$

with a very poor frequency localization.

To summarize, the PD was found to be the only useful TFR of the ones considered for the identification of business cycle. The business cycle period that is estimated from the US financial data is in the range of 3 to 4 years as reported in some previous studies –see, for example, [Chen (1996a)]. Therefore, our

estimation of 3.85 years for the business cycle period estimated from the NASDAQ index is in line with the previous estimates, demonstrating the robustness of our results and providing additional evidence in support of the usefulness of the PD.

5. Conclusions

This paper aimed to illustrate the potential of some of the TFR techniques widely used in various fields of science and engineering for use in the analysis of stock market data and to compare their performances. For this purpose, we used two sets of stock market data and analyzed their behavior using two linear (the Gabor Transformation and the Short Time Fourier Transform) and two quadratic (the Wigner Distribution and the Page Distribution) TFRs. The ISE-100 index of Istanbul Stock Exchange over the period from July 9, 1990 to December 25, 2000 was picked to perform the analysis on the data from an emerging market, whereas the NASDAQ-100 series (made up of index values recorded from November 14, 1985 to March 13, 2001) was used to evaluate the performance of the TFR techniques considered within the context of a well-established market.

For comparative performance evaluations of these TFRs in detecting and decoding cycles that may be present in the original ISE and NASDAQ data, we utilized a specially synthesized time series whose trend and/or cycle components can be analytically specified and computed. We constructed this series in such a way to mimic the pattern of a typical stock index series and let it serve as a benchmark in our simulated performance analysis.

The results we obtained using the synthesized index revealed that the WD mainly suffers from the ambiguity caused by cross terms, while the STFT and GT results exhibit serious problems in frequency localization. Consequently, the WD, STFT and GT could identify only one or two cycles out of the three inserted into the series, and in an extremely ambiguous fashion at times. The PD, on the other hand, performed impressively well, detecting all of the cycles known to exist and estimated their respective periods quite accurately. The performance of the PD remained satisfactory even after a realistic amount of white Gaussian noise was added to the synthesized index.

The results obtained from the actual ISE-100 and NASDAQ-100 indices indicated that analyses based on the GT and STFT may be useful to identify the business cycle but the WD analysis gives no clear indication of the presence of a business cycle. Furthermore, the results from the GT and the STFT applications raise serious concerns about the accuracy of the estimated cycle period, as they seriously suffer from extremely poor frequency localization. The cycle identification results produced by the PD, on the other hand, turn out to be very successful in estimating the business cycle period with a high confidence.

This superior performance of the Page distribution, used for the first time in economics/finance literature, makes it a very promising tool for use in these fields.

Finally, even though the relatively short data recording periods might restrict the frequency resolution of the TFR analysis, increased use of such techniques in the analysis of economic and financial time series with longer time frames and moderately large sample sizes would be a welcome addition to the literature.

References

Alper, C. E. (2000). *Stylized Facts of Business Cycles, Excess Volatility and Capital Flows: Evidence from Mexico and Turkey*. Bogazici University Center for Economics and Econometrics Working Paper No. ISS/EC-00-07, Istanbul.

Bjornland, H.C. (2000). Detrending Methods and Stylized Facts of Business Cycles in Norway: An International Comparison. *Empirical Economics*, 25(3). pp. 369-92.

Bowden, R.J. and V.L. Martin (1993). Reference Cycles in the Time and Frequency Domains: Duality Aspects of the Business Cycle. In P.C.B. Phillips (ed.), *Models, Methods, and Applications of Econometrics: Essays in Honor of A. R. Bergstrom*, Cambridge, MA and Oxford: Blackwell. pp. 201-19.

Chen, P. (1996a). A Random Walk or Color Chaos on the Stock Market? Time-Frequency Analysis of S&P Indexes. *Studies in Nonlinear Dynamics and Econometrics*, 1(2). pp. 87-103.

Chen, P. (1996b). Trends, Shocks, Persistent Cycles in Evolving Economy: Business-Cycle Measurement in Time-Frequency Representation. In W.A. Barnett, A.P. Kirman and M. Salmon (eds.), *Nonlinear Dynamics and Economics: Proceedings of the Tenth International Symposium in Economic Theory and Econometrics (International Symposia in Economic Theory and Econometrics Series)*, Cambridge, New York and Melbourne: Cambridge University Press. pp. 307-31.

Cohen, L. (1995). *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice Hall, 1995.

Fama, E.F. and K.R. French (1988). Permanent and Temporary Components of Stock Prices. *Journal of Political Economy*, 96(2). pp. 246-73.

Gabor, D. (1946). Theory of Communication. *Journal of the Institute of Electrical Engineers (London)*, 93(3). Part I, pp. 429-457.

Garcia-Ferrer, A. and R.A. Queralt (1998). Using Long-, Medium-, and Short-Term Trends to Forecast Turning Points in the Business Cycle: Some International Evidence. *Studies in Nonlinear Dynamics and Econometrics*, 3(2). pp. 79-105.

Greenblatt, S. A. (1997). Wavelet Basis Selection for Regression by Cross-Validation. In H. Amman, B. Rustem and A. Whinston (eds.), *Computational*

Approaches to Economic Problems (Advances in Computational Economics, vol. 6), Dordrecht, Boston and London: Kluwer Academic, 1997. pp. 39-55.

Hlawatsch, F. and G.F. Boudreaux-Bartels (1992). Linear and Quadratic Time-Frequency Signal Representations. *IEEE Signal Processing Magazine*, 9(2). pp.21-67.

Hodrick, R. J., and E. C. Prescott (1997). Post-War U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking*, 29(1). pp. 1-16.

Hong, Y. (1999). Hypothesis Testing in Time Series via the Empirical Characteristic Function: A Generalized Spectral Density Approach. *Journal of the American Statistical Association*, 94(448). pp. 1201-20.

Jacquier, E. and V. Nanda (1991). *Cyclical Components in Stock Market Returns*. Manuscript, Ithaca, NY: Cornell University.

King, R.G. and S.T. Rebelo (1993). Low Frequency Filtering and Real Business Cycles. *Journal of Economic Dynamics and Control*, 17(1-2). pp. 207-31.

Lee, G.H. (1998). Wavelets and Wavelet Estimation: A Review. *Journal of Economic Theory and Econometrics*, 4(1). pp. 123-57.

Levy, D. and H. Chen (1994). Estimates of the Aggregate Quarterly Capital Stock for the Post-war U.S. Economy. *Review of Income and Wealth*, 40(3). pp. 317-49.

Lo, A. and C. MacKinlay (1988). Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test. *Review of Financial Studies*, 1. pp. 41-66.

Metin-Ozcan, K., E. Voyvoda and E. Yeldan (2001). Dynamics of Macroeconomic Adjustment in a Globalized Developing Economy: Growth, Accumulation and Distribution, Turkey 1969-1998. *Canadian Journal of Development Studies*, 22 (1). pp. 219-253.

Poterba, J. and L. Summers (1988). Mean Reversion in Stock Prices: Evidence and Implications. *Journal of Financial Economics*, 22. pp. 27-59.

Qian, S., and D. Chen (1996). *Joint Time-Frequency Analysis*. Upper Saddle River, NJ: Prentice-Hall.

Qian, S. and D. Chen (1993). Discrete Gabor Transform. *IEEE Transactions on Signal Processing*, 41(7). pp. 2429-2438.

Sargent, T. (1979). *Macroeconomic Theory*. New York, San Fransisco and London: Academic Press.

Thoma, M.A. (1994). The Effects of Money Growth on Inflation and Interest Rates across Spectral Frequency Bands. *Journal of Money, Credit and Banking*, 26(2). pp. 218-31.

Turhan-Sayan, G. and M. Karaduman (2001). Use of PCA and Quadratic TFR Techniques in Electromagnetic Target Classification from Scattered Data. Paper Presented at the *IEEE AP-S International Symposium and USNC/URSI National Radio Science Meeting*, Boston, MA, July 8-13.

Turhan-Sayan, G. and S. Sayan (2001). A Comparative Evaluation of the Performances of Different Filtering Techniques in Business Cycle Identification. Paper Presented at the *Seventh Annual Meetings of the Society for Computational Economics*, New Haven, CT, June 28-30.